BIBLIOTEGA DI SC. STATISTICHE DEM. E ATTUARIALI
SERVIZIO BIBLIOTEGARIO NAZIONALE
BID 882410 CID
ACO. 981 / 102 INV. 82765
COLL 5-COP WP-2002/16

Accounting for heavy tails in stochastic frontier models

A. Tancredi

2002.16

Dipartimento di Scienze Statistiche Università degli Studi Via C. Battisti 241-243 35121 Padova 235 N 20 9W 123 7

Accounting for heavy tails in stochastic frontier models

iberemal A

2002.16

Dipartimento di Scienze Statistiche Università degli Studi Via C. Bartisti 241-143 35121 Padova

# Accounting for Heavy Tails in Stochastic Frontier Models

Andrea Tancredi \*
Dipartimento di Scienze Statistiche
Università di Padova

#### Summary

This paper aims at introducing a new class of stochastic frontier models that can take account for fat tails in the composed error. Quite surprisingly, all the stochastic frontier models proposed in literature cannot handle situations where the empirical distribution of the composed error has heavy tails. These situations are instead very common in applications. In particular, we will propose to model the composed error with the skew-t distribution. This is equivalent to assume a Student-t distribution for the measurement error and a half-t distribution for the inefficiency. In this way, we extend quite naturally, the stochastic frontier model where a normal distribution is assumed for the symmetric error and a half-normal distribution is assumed for the inefficiency term.

Some key words: Composed error; Efficiency analysis; Skew-t distribution

<sup>\*</sup>Dipartimento di Scienze Statistiche, Università di Padova, Via Cesare Battisti 241-249, 35121, Padova, Italy; E-mail tancredi@stat.unipd.it

# Accounting for Heavy Balls in Stochastic Trentier Violets

Andrea de Seu edi. Andreamp de Seu edine edine edile Europe de Antre de Constant

#### veisione a fil

All the common the common the company of the management of the company of the com

#### 1 Introduction

Stochastic frontier models, introduced in Meeusen and van den Broeck (1977) and Aigner et. al. (1977), are useful tools to evaluate the efficiency of economic agents, such as firms, individuals or countries. The principle underlying the model is that the observed production of a single unit cannot exceed the unobserved potential production i.e. the frontier, which is the maximum possible production given input quantities. The difference between the frontier and the observed production is a measure of inefficiency. Such difference is modelled by a one-sided random variable. The model is completed adding a symmetric random variable capturing the mesaurement error of the frontier.

In the commonly adopted formulation a normal distribution is assumed for the symmetric error while a half-normal distribution is assumed for the one-sided random variable. This framework leads to model the composed error by a probability distribution known, in the statistical literature, as skew-normal distribution, see Azzalini (1985) and Azzalini and Capitanio (1999). The skew-normal distribution is a family of distributions including the normal, but with an extra parameter to regulate the skewness. The stochastic frontier model introduced by Aigner et. al (1977), ALS model in the following, is a skew normal regression models with negative skewness.

Several other distributions for the inefficiency term, have been proposed, in different times, in place of the half-normal distribution. For example, Meeusen and van den Broeck (1977) adopt the exponential distribution, Stevenson (1980) the truncated normal and Greene (1990) the gamma distribution. A unified approach is proposed in van den Broeck et al (1994) where the results obtained with different inefficiency distributions are pooled together by Bayesian model averaging. Finally, a semi-parametric Bayesian approach is proposed in Griffin and Steel (2002) where a Dirichelet process with gamma mean is assumed for the inefficiency distribution.

This paper aims at introducing a new class of stochastic frontier models that can take account for fat tails in the composed error. Quite surprisingly, all the proposals existing in literature, at least to our knowledge, cannot handle situations where the empirical distribution of the composed error has heavy tails which, instead, are very common in applications. In particular, we will propose to model the composed error with the skew-t distribution introduced by Azzalini and Capitanio (2002). This is equivalent to assume a Student-t distribution for the measurement error and a half-t distribution for the inefficiency. In this way, we extend quite naturally, the ALS model, which becomes a limit case of our model.

Section 2 describes the new model with particular emphasis on frequentist inference both for testing the presence of the inefficiency term and for estimating individual technical efficiencies. In particular it is shown that, in estimating individual efficiencies, the skew-t model has a completely different approach respect to the ALS model when we have observations which are suspected of being outliers. In the ALS model, observations with a large postive deviation from the estimated frontier lead to estimates of individual efficiency concentrated in one. Instead, for the skew-t model these observations are not considered informative for estimating individual efficiency.

In section 3 we apply the model to the well known data set of the American elettrical companies. This data set has been carefully analyzed, in a frequentist setting, by Ritter and Simar (1994), who conclude that the data show more evidence for the normal linear model without inefficiency than for stochastic frontier models. We will show that assuming a skew-t model for the composed error provides a more reasonable fit than the normal linear model and that taking account for fat tails increases the evidence for the inefficiency term. Moreover we will study the behaviour of the estimates of individual efficiencies respect to the distance from the estimated frontier both for the ALS model and the skew-t model.

In section 4 we give a brief discussion for subsequent modelling and research.

As the single load of the second and make how gained the state of a state of selections of the control of the state of the selection of the se

particular de la company d La company de la company d La company de la c

This paper aims at introducing a new cines of stockness. Fromiss, modesschat can take acoung for rate took proposed scan alignite empressingly, all the proposed existing the literatures at lower to ear to which instead, are very common in applications. In particular, which instead, are very common in applications. In particular, which instead, are very common in applications. In particular, which instead, are very common in applications to particular, which is account a start of distribution into the composed error with the short distribution into the contract of account a light of the distribution of the light contract of the which which alternating the ACS model, which

Social Street the street of the mewerhood with particular employer or broadening antiference both for particular employers to the confidence of the sating that vidential particular and particular to show that the confidence of the sating that vidential efficient as the sating of th

In social Same apply the middeless the well knowned a set of the Ameden ear script at Companies of the dame set in a local to a test of the set of the dame set in a frequential set of the date and Singe (1994), who could also the date show more evidence for the providing a slower model without includency than the composed error provides a more reasonable me than the mental linear model and time taking account for factor measured error measured time taking account for factor measured and time taking account for factor measured as a vice with said of the behaviour of the same at individual access respect to the distance true the estimated from the behaviour for the

the second 4 we give a brief discussion for subsequent modelling and research

## 2 The model

Let us consider the model

$$y_i = h(x_i, \beta) + \epsilon_i - z_i \qquad i = 1, \dots, n$$
 (1)

where  $y_i$  denotes the log of the output variable for firm i (i = 1, ..., n) and  $x_i$  is a vector of observations for the explanatory variables for firm i. In stochastic frontier models  $\epsilon_i$  is a symmetric distribution with zero mean while  $z_i$  is a one-sided positive distribution. For istance, in the ALS model  $\epsilon_i$  is assumed  $\mathcal{N}(0, \sigma_\epsilon^2)$  and  $u_i$  is assumed half-normal distributed  $|\mathcal{N}(0, \sigma_z^2)|$ .

Respect to the ALS model we assume that  $z_i = |v_i|$  and that the couples  $\epsilon_i, v_i$  are distributed like a bivariate Student-t distribution with zero means, scale parameters  $\sigma_{\epsilon}$  and  $\sigma_z$ , uncorrelated components and shape parameter  $\nu$ , independently for i = 1, ..., n. The density can be written as

$$f(\epsilon_i, v_i) = \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \nu \pi \sigma_{\epsilon} \sigma_z} \left\{ 1 + \frac{1}{\nu} \left[ \left(\frac{\epsilon_i}{\sigma_{\epsilon}}\right)^2 + \left(\frac{v_i}{\sigma_z}\right)^2 \right] \right\}^{-\frac{\nu+2}{2}} \qquad i = 1, \dots, n.$$
 (2)

The meaning of the variables  $\epsilon_i$  and  $z_i$  remains the same of standard stochastic frontier models. In fact  $\epsilon_i$  still represents a symmetric disturbance capturing the measurement error of the stochastic frontier and  $z_i$  is still a nonnegative random variable modelling the level of inefficiency. The novelty here is that  $\epsilon_i$  is marginally distributed like a univariate Student-t distribution  $\mathcal{T}(0, \sigma_{\epsilon}, \nu)$  and  $z_i$  is marginally distributed like a half Student-t  $|\mathcal{T}(0, \sigma_z, \nu)|$ .

The sampling distribution of  $y_i$  has been derived, in a more general context, by Azzalini and Capitanio (2002). Specifically, if we write

$$y_i = h(x_i, \beta) + \sqrt{\sigma_{\epsilon}^2 + \sigma_z^2} \left[ \frac{\sigma_{\epsilon}}{\sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}} U - \frac{\sigma_z}{\sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}} |U_0| \right], \tag{3}$$

where  $(U_0, U)$  is a standard bivariate Student-t distribution with shape parameter  $\nu$ , then we can apply proposition 9 of Azzalini and Capitanio (2002). Thus we have that

$$p(y_i; \beta, \sigma_{\epsilon}, \sigma_z, \nu) = 2 f_t(y_i; h(x_i, \beta), \omega, \nu)$$

$$\times T_1 \left[ \alpha \frac{y_i - h(x_i, \beta)}{\omega} \left( \frac{\nu + 1}{\omega^{-2} (y_i - h(x_i, \beta))^2 + \nu} \right)^{1/2}; \nu + 1 \right]$$

$$(4)$$

where  $f_t(y; h(x_i, \beta), \omega, \nu)$  denotes the density function of a Student-t distribution with mean  $h(x_i, \beta)$ , scale  $\omega$  and  $\nu$  degrees of freedom,  $\omega = \sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}$ ,  $T_1(y, \nu + 1)$  denotes the scalar Student-t distribution with  $\nu + 1$  degrees of freedom and  $\alpha = -\sigma_z/\sigma_{\epsilon}$ . Distributions with density (4) are called skew-t distributions and they generalize the skew-normal distributions (Azzalini, (1985)) which can be obtained when the shape parameter  $\nu$  goes to infinity.

Note that the same mechanism that generates a Student-t distribution from a normal distribution allow us to generate a skew-t from a skew-normal. In fact Azzalini and Capitanio (2002) show that a skew-t distribution can be obtained as mixture of skew-normal variates with scale parameter  $1/\sqrt{\lambda}$  where  $\lambda \sim \Gamma(\nu/2, \nu/2)$ . Thus model (1) can be written also in the following way

$$y_i = h(x_i, \beta) + \frac{1}{\sqrt{\lambda}} (\epsilon_i - z_i) \qquad i = 1, \dots, n$$
 (5)

where  $\epsilon_i \sim \mathcal{N}(0, \sigma_{\epsilon}^2)$ ,  $z_i = |\mathcal{N}(0, \sigma_z^2)|$  and  $\lambda \sim \Gamma(\nu/2, \nu/2)$ . Indeed, the difference  $\epsilon_i - z_i$  in (5), which is the composed error of the ALS model, follows a skew-normal distribution with location parameter equal to zero, scale parameter equal to  $\sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}$  and shape parameter equal to  $-\sigma_z/\sigma_{\epsilon}$ .

(I) I was the second of the contract of the co

to control to such an horal (in the first of the most of the control of the gold of the characters approached a promotive and the control of the control of the second and the control of the second and the control of the control of

"I be a complete want if the exact of constant of exact of exact or own in the least of the exact of exact of the constant of

e sang dag distribution of y that bein decimal to compare, negation of a distribution of a superior distribution

where (fig. '.' is a standard bysniala bindens-t distribution of pringer phrodesses of movered or hippy problems for 9 of Asigh and Coplanto (Mull). This ye love that

where  $\mu_{ij}$  and the colour of the decided and the colour of the colour field  $\mu_{ij}$  and  $\mu_{ij}$  are  $\mu_{ij}$  and  $\mu_{ij}$  and the field  $\mu_{ij}$  and the same and they galaxies the second distribution of the colour  $\mu_{ij}$  and  $\mu_{ij}$  and the second colour  $\mu_{ij}$  and the second distribution of the second  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{ij}$  and  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{ij}$  and  $\mu_{ij}$  are the second  $\mu_{ij}$  are the second  $\mu_{ij}$  and  $\mu_{$ 

Note that the same adjusted of the State of

$$\mathbf{r}_1, \dots, \mathbf{r}_{n-1} \in \{(s_n, s_n) \mid s_n \in \mathbb{N} \}$$

scheep of the Charles of the CV (tree) had A self (1922) and A self (1922) and the composed and the composed and the CV so will be composed and the composed and the CV so middle collowers a land of the mail the continuous visit of the control of

## 2.1 Testing the presence of the inefficiency term

The natural use of model (4) is whitout constraining the parameter  $\alpha$  to be negative. In fact, if we let  $\alpha$  to vary in  $(-\infty, \infty)$  the statistical model (4) is a regression model that can account for fat tails and for both positive and negative skewness of the error distribution. Positive values for  $\alpha$  correspond to composed error distributions where a half-t is added to a Student-t, switching drastically the meaning of the model respect to stochastic frontier models. Anyway considering the model (4) with  $\alpha \in (-\infty, \infty)$  can be helpful for inferential aims also in stochastic frontier models. Specifically, in testing  $H_0: \alpha = 0$  against  $H_1: \alpha < 0$ , i.e. in testing the presence of the inefficiency term, the signed version of the likelihood ratio statistic may be used. That is

$$R = sgn(\hat{\alpha}) \{ 2[\ell(\hat{\theta}) - \ell(\hat{\theta}^*)] \}^{1/2}$$
(6)

where  $\theta = (\beta, \omega, \alpha)$ ,  $\hat{\alpha}$  and  $\hat{\theta}$  denote the maximum likehood estimates of  $\alpha$  and  $\theta$  when  $\alpha \in (-\infty, \infty)$  and  $\hat{\theta}^*$  denote the maximum likelihood estimate when  $\alpha = 0$ . The asymptotic distribution of R under the null model is standard normal and having observed  $R_{obs}$  the evidence of  $H_0$  against  $H_1$  is given by  $P(\mathcal{N}(0,1) < R_{obs})$ .

#### 2.2 Estimation of firm-level technical efficiencies

In stochastic frontier models the main interst is not on the parameters themselves, but in the individual technical efficiencies, measured by  $r_i = \exp(-z_i)$ . Estimates of these quantities are obtained considering the conditional expected values  $E(r_i|y_i)$ , see for example Coelli *et al.* (1998). In the appendix we prove that for model (1) the conditional density of  $z_i$  given  $y_i$  is

$$f(z_i \mid y_i) = \frac{\left(\frac{\nu}{2}\right)^{\nu/2} \Gamma\left(\frac{\nu+2}{2}\right)}{f(y_i) \Gamma\left(\frac{\nu}{2}\right) \pi \sqrt{\sigma_{\epsilon}^2 \sigma_z^2} \left[\frac{1}{2} \left\{\nu + \frac{(y_i - h_i)^2}{\sigma_{\epsilon}^2 + \sigma_z^2} + \frac{\sigma_{\epsilon}^2 + \sigma_z^2}{\sigma_{\epsilon}^2 \sigma_z^2} \left(z_{\mid} + (y_i - h_i) \frac{\sigma_z^2}{\sigma_{\epsilon}^2 + \sigma_z^2}\right)^2\right\}\right]^{\frac{\nu+2}{2}}}.$$
 (7)

Thus, a point estimation of the individual technical efficiency can be obtained integrating numerically  $\int e^{-z_i} f(z_i|y_i) dz_i$  where in  $f(z_i|y_i)$  we replace the unknown parameters with the maximum likelihood estimates. A measure of uncertainty of these estimates is obtained considering the plugin estimates of the standard deviation of  $e^{-z_i}$  given  $y_i$ . Note that, in this ways, we not take account of parameter uncertainty, but, from a frequentist point of view, this seems the standard practice since Jondrow  $et\ al.\ (1982)$ .

Let us observe that adopting the skew-t model in place of the ALS model lead to a completely different behaviour in estimating individual efficiency of firms with large positive deviations from the estimated frontier. In fact, if we indicate the frontier  $h(x_i, \beta)$  with  $h_i$  we have that (see the appendix) for the ALS model

$$\lim_{y_i - h_i \to \infty} f(z_i | y_i) = \begin{cases} \infty & if \quad z_i = 0\\ 0 & if \quad z_i > 0 \end{cases}$$
 (8)

while for the skew-t model

$$\lim_{y_i - h_i \to \infty} f(z_i | y_i) = 0 \qquad z_i \ge 0.$$
(9)

This means that when  $y_i - h_i$  goes to infinity the conditional distribution of  $z_i$  in the ALS model is concentrated in zero while for the skew-t model is improper uniform on  $[0, \infty)$ .

Now let us suppose to have an observation  $y_i$  which produces an estimated positive residual  $y_i - h(x_i, \hat{\beta})$  very far from the bulk of the other residuals. Thus we should consider  $y_i$  as an

Resting the presence of the medicioner term

ti provide de conseque en arte en elemente est generalement parador el projetifico en formarer esta del fina con una viola labora morganica ir el (e) l'elementation besolt (como el fini e avent es nel socialiste de la fini e avent es nel socialiste de la contra de la como el fini e avent en l'appendinte de la company de

1 (1) (19) 3 to (10) 1 (10) 1 (10) 1 (10)

o postantina de la composición de la c La composición de la La composición de la

- Not agnivithed to regional forest-possible constructions; in the

t tradición de la medición de tegenda per en el secono como de como en encolor y incidencia de la medición y m Procedida del persona de la companiente de la companiente de la companiente de la companiente de la companient Professione de la companiente del companiente de la companiente del companiente de la companiente del companiente de la companiente del companiente de la companiente del companie

erang a pear comment on all the antivided pedation of ficiency can be obtained interpetion from period of the promise featign of figure of the relation of the

wille for the slow-t model

 $\hat{\omega} = \hat{\omega} \leq \hat{\omega} = 0$  (i.e.  $\hat{\omega} = \hat{\omega} = \hat{\omega}$ 

Istore 2.1A ods in 32.10 nicialid tutid Langitihares 910 ydnesnjih seny 10 - 32 nestve itejs v semeserar Loo 01 to corolline regionjani zi lahom 3-vests nici sel silištv corollini test capanitiseji landese voticeg besamitise na regioni didir je nestve sede nes visit od izanime se sede nicialismi se visit od nicialismi v odinares bloods sve sed 1 - sladbilar meto sede to dom se most me visit Associalismi se se se outlier and, maybe, we should remove it from the data before calculating the required inference. Now, if we are fitting the data with the ALS model we have, for this observation, that the plug-in estimate of  $r_i$  will be near one and the plug-in estimates of  $Var(e^{-z_i}|y_i)$  near zero. In fact, by (8) the distribution of  $e^{-z_i}$  given  $y_i$  will be very concentrated near one. Thus we should believe that this observation is very informative for estimating individual technical efficiency. The situation is completely different with the skew-t model. In fact when we have an observation which produces a high estimated residual, the plug-in estimate of  $Var(e^{-z_i}|y_i)$  will be very high. This is because the limit distribution of  $z_i|y_i$  is completely flat. Thus, this observation will not considered very informative for estimating individual technical efficiency. This means that, when our primary goal is to estimate individual efficiencies, by adopting the skew-t model we do not have to worry whether to discard an outlier or not. In fact the model automatically increases the degree of uncertainty of our conclusions when we are in presence of outliers.

## 3 Example

We consider the data collected by Christensen and Greene (1976) for 123 electric utility companies in the US in 1970. The data are given in the appendix to Greene (1990) and have been used by van den Brock et al (1994) and Tsionas (2002). There are three production factors labor, capital and fuel with prices  $p_L$ ,  $p_K$ , and  $p_F$  and the cost function which is usually specified is

$$y_i = -\beta_0 - \beta_1 \ln Q_i - \beta_2 \ln^2 Q_i - \beta_3 \ln \frac{p_{K_i}}{p_{F_i}} - \beta_4 \ln \frac{p_{L_i}}{p_{F_i}} + \epsilon_i - v_i$$
(10)

where  $y_i = -\ln C_i/p_{F_i}$ ,  $Q_i$  is the output and  $C_i$  the cost of the *ith* firm.

## 3.1 Normal against skew-t model

For this data set, Ritter and Simar (1994) compare, from a likelihood point of view, a normal regression model with several stochastic frontier models. They conclude that a normal linear model without inefficiency is enough to explain the data. Thus we firstly try to understand if a skew-t model lead to a significantly improved explanation of the data over a standard regression model.

Maximum likelihood estimates (MLE) and approximated 95% confidence interval are reported in Tab 1 for the normal linear model, the ALS model and the skew-t regression model. We observe that the maximum log-likelihood for the normal model is 65.67, while for the skew-t regression model we have obtained 68.75. Thus the test statistic  $D = \ell(\hat{\theta}) - \ell(\hat{\theta}^*)$ , where  $\ell(\hat{\theta})$  and  $\ell(\hat{\theta}^*)$  denote the maximized log-likelihood within the skew-t regression model and the normal regression model, is 6.12. Anyway, comparing the two models by the test statistic D needs some caution. In fact the normal regression model occurs on the frontier of the parametric space of the skew-t regression model, specifically when  $\nu$  tends to infinity and the ratio  $\sigma_z/\sigma_\epsilon$  tends to zero. Therefore, the chisquare approximation to twice the log-likelihood does not apply in this context. To circumvent the problem we have opted for a bootstrap approach. Specifically we have generated 10000 samples from the estimated normal model  $y_i = h(x_i, \beta) + \epsilon_i$  with  $\epsilon_i \sim \mathcal{N}(0, \hat{\sigma})$  where  $\beta$  and  $\hat{\sigma}$  are the MLE. For each sample we have calculated the difference between the maximum log-likelihood under the skew-t model with  $\nu$  fixed to the MLE  $\hat{\nu}=4.6461$  and the maximum log-likelihood under the normal model. The associated bootstrap p-value for the normal model is 0.014. This indicates low evidence for the normal linear model respect to the skew-t model. Graphical analysis seem to confirm this. In Fig 1 we report, for the normal, ALS, and skew-t model, the histograms and pp-plots for the estimated standardized residuals  $(y_i - h(x_i, \hat{\beta}))/\hat{\omega}$ , where  $\omega$  is equal to  $\sigma_{\epsilon}$  for the normal model and to  $\sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}$  both for the ALS model and the skew-t model. In particular,

omited and mayber we should remove in items the standard contribution of the representation of the contribution of the contribution.

# elepergy: (i.

e in againe of the simbole and the COOL energy for agentificational industrial production of the confidence of The agentification (Place) and the energy of the energy for the confidence of the confidence of the energy for the confidence of the co

which  $y \in \{-1\}$  is  $\mathbb{C}_{\mathbb{Z}^2(y)} (z, t)$  is the exagnic and  $\mathbb{C}_{\mathbb{Z}}$  the cost of the z / t firm.

## 3.1 Formal against skew-t model

For the data set filter and diesa (1994) compare, from a life-libeod pent of view, a normal even vion model villi event stockers, fromier models. They conclude that a normal through model without inclinated or are enough as explainable data. Thus we limbly my to nadorstand data slees anded to data algunisquate arquerse explanation of the data over a standard represent

May can highline direction to the ALE mudal and 25% confidence discrete to reported in Tab 1 for the name of the normal linear model, its ALE mudal and the slows to restrict the reason of the restrict the reason of the name of the slows to the restrict the restrict that the name of the name of the slows to the restrict the slows to the restrict the name of the

looking the pp-plots for the normal model and the skew-t model it is evident that the latter fits the data better.

#### 3.2 Testing the inefficiency term

The next goal in analazying the data is to understand if we can drop the inefficiency term when we take account of heavy tails, thus we test  $H_0: \alpha=0$  against  $H_1: \alpha<0$ , in model (4) with  $\alpha\in(-\infty,\infty)$  and we use test statistic R. The MLE of  $\alpha$  is  $\hat{\alpha}=-0.936$  and the maximum log-likelihood associated to the Student-t regression model is 67.56. The test statistic  $R=sign(\hat{\alpha})(\ell(\hat{\theta})-\ell(\hat{\theta}^*))^{1/2}$  is -1.536 and the associated observed p-value is 0.06 indicating low evidence for the absence of the inefficiency term. Note that inference on the inefficiency term is affected by accounting for heavy tails. Indeed, in comparing the normal model against the ALS model, we may test  $H_0: \alpha=0$  against  $H_1: \alpha<0$ , in model (4) with  $\alpha\in(-\infty,\infty)$  and  $\nu=\infty$ . In this case the test statistic R is equal to -0.94 and the observed p-value is  $P(\mathcal{N}(0,1)<-0.94)=0.17$ . Thus, for the elettric companies data set, when we allow for thick tails the evidence for the presence of the inefficiency term increases.

## 3.3 Estimating individual efficencies

Finally we present our results for the efficiency of firms within the sample. In table 2 we compare the quantities  $r_i = E(e^{-z_i}|y_i)$  for the first five firms in the sample both for the ALS model and the skew-t model. These are the same firms analyzed by van den Broeck et~al.~(1994). We see that, with the skew-t model the estimates for  $r_i$  are generally bigger than for the ALS model, with the exception of the second firm. More insight into the beheaviour of the models about efficiency analysis is given in Fig 2. This plots, for all the firms, the plug-in estimates of  $r_i = E(e^{-z_i}|y_i)$  and  $s_i$  where  $s_i^2 = Var(e^{-z_i}|y_i)$  against the estimated residuals  $y_i - h(x_i, \hat{\beta})$  both for the ALS model and the skew-t model. We see that for the ALS model  $r_i$  is always increasing while  $s_i$  first increases and then decreases in according to the fact the  $e^{-z_i}$  given  $y_i$  should be concentrated in one when  $y_i - h(x_i, \hat{\beta})$  becomes very large. For the skew-t model we have a different behaviour, when the residuals becomes larger both  $r_i$  and  $s_i$  start to increase in according with our findings that the density of  $z_i$  given  $y_i$  is completely flat when  $y_i - h(x_i, \hat{\beta})$  becomes very large.

#### 4 Discussion

In this paper we have introduced a new stochastic frontier model with several attractive features. First of all, it generalizes the common stochastic frontier model, where a normal distribution is assumed for the error term and a half-normal distribution is assumed for the inefficiency term, allowing for fat tails in the composed error distribution. Adopting this new model we do not have to worry to remove outlier observations before drawing inference. In fact, if our aim is to estimate individual technical efficieny, the model automatically increases the uncertainty of our estimates when we have observations lying above and far from the estimated frontier. We have discussed frequentist inference and further research will be conducted on estimating the model from a Bayesian point of view.

locking this pp. planstor the remainments and this show. "Goodel is brevident that the latter its The term better."

#### med geneicitani s.P. getroff - f. E

tes or much so colling the grant was our internal of the out of the principal deciment in a property of the solution of the so

#### asistemblica state dibai pojeno alekil

construction of the response to the statement of the second construction of the construction and construction of the construction and the construction of the construction of the construction of the construction of the second construction

#### recipelate ATT

in this paper so have monthered a new stochastic frontier model with several attraction features. First of all at generalizes the common stochastic frontier model, where a control of religionary safetying for the error term and a half-derival distribution. Adopting this new mode each conception allowing for the religion of the characteristic framework to the characteristic observations before drawing independent. It is the control of our estimates when we have observations igning allow, and for how estimated from our to the continued from the design of our drawing there are said for the continued of our drawing observations ignificant allows and the continued of our estimated from our the model for a discussion of our discussion of the continued for all the continued of the con

# Appendix A

Density for the individidual inefficiency in the skew-t model

$$f(z_i \,|\, y_i) = \int f(z_i \,|\, \lambda_i, y_i) f(\lambda_i | y_i) d\lambda_i$$

$$f(z_i \mid \lambda_i, y_i) = \frac{e^{-\frac{1}{2} \frac{\lambda_i (\sigma_{\epsilon}^2 + \sigma_z^2)}{\sigma_{\epsilon}^2 \sigma_z^2} \left( z_i - \frac{-\sigma_z^2 (y_i - h_i)}{\sigma_{\epsilon}^2 + \sigma_z^2} \right)^2}}{\sqrt{2\pi \frac{\sigma_{\epsilon}^2 \sigma_z^2}{\lambda_i (\sigma_{\epsilon}^2 + \sigma_z^2)}} \Phi\left( \frac{-\sqrt{\lambda_i} \sigma_z (y_i - h_i)}{\sigma_{\epsilon} \sqrt{\sigma_{\epsilon}^2 + \sigma_z^2}} \right)}$$

$$f(\lambda_i|y_i) = \frac{f(y_i|\lambda_i)f(\lambda_i)}{f(y_i)} = \frac{2\sqrt{\lambda_i}e^{-\frac{1}{2}\frac{\lambda_i(y_i-h_i)^2}{\sigma_z^2+\sigma_\epsilon^2}}\Phi\left(\frac{-\sqrt{\lambda_i}\sigma_z(y_i-h_i)}{\sigma_\epsilon\sqrt{\sigma_\epsilon^2+\sigma_z^2}}\right)}{f(y_i)\sqrt{2\pi(\sigma_\epsilon^2+\sigma_z^2)}}\frac{\left(\frac{\nu}{2}\right)^{\nu/2}e^{-\lambda_i\frac{\nu}{2}}\lambda_i^{\nu/2-1}}{\Gamma\left(\frac{\nu}{2}\right)}$$

$$f(z_{i} | y_{i}) = \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{f(y_{i})\Gamma\left(\frac{\nu}{2}\right)\pi\sqrt{\sigma_{\epsilon}^{2}\sigma_{z}^{2}}} \int_{0}^{\infty} \lambda_{i}^{\frac{\nu+2}{2}-1} e^{\frac{-\lambda_{i}}{2}\left[\nu + \frac{(y_{i}-h_{i})^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}} + \frac{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}\sigma_{z}^{2}}\left(z_{i} + (y_{i}-h_{i})\frac{\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}\right)^{2}\right]} d\lambda_{i}$$

$$= \frac{\left(\frac{\nu}{2}\right)^{\nu/2}\Gamma\left(\frac{\nu+2}{2}\right)}{f(y_{i})\Gamma\left(\frac{\nu}{2}\right)\pi\sqrt{\sigma_{\epsilon}^{2}\sigma_{z}^{2}}\left[\frac{1}{2}\left\{\nu + \frac{(y_{i}-h_{i})^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}} + \frac{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}\sigma_{z}^{2}}\left(z_{i} + (y_{i}-h_{i})\frac{\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}\right)^{2}\right\}\right]^{\frac{\nu+2}{2}}}$$

# Appendix B

Limiting behaviour for the individual inefficiency  $1)\ ALS\ model$ 

Let us consider the following notation  $A = \frac{1}{2} \frac{\sigma_z^2 \sigma_\epsilon^2}{\sigma_z^2 + \sigma_\epsilon^2}$ ,  $B = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_z^2}$ ,  $C = \frac{\sigma_z}{\sigma_\epsilon} \frac{1}{\sqrt{\sigma_\epsilon^2 + \sigma_z^2}}$ , and observe that  $-\frac{1}{2}C^2 = -AB^2$ . We have

$$\lim_{y_i - h_i \to \infty} f(z_i | y_i) = \lim_{y_i - h_i \to \infty} \frac{e^{-A(z_i + B(y_i - h_i))^2}}{\Phi(-C(y_i - h_i))} = \lim_{y_i - h_i \to \infty} \frac{-2e^{-A(z_i + B(y_i - h_i))^2} AB(z_i + B(y_i - h_i))}{-Ce^{-\frac{1}{2}C^2(y_i - h_i)^2}}$$

Thus if  $z_i = 0$  we have

$$\lim_{y_i - h_i \to \infty} f(z_i | y_i) = \lim_{y_i - h_i \to \infty} 2 \frac{AB^2}{C} (y_i - h_i) = \infty$$

while if  $z_i > 0$  we have

$$\lim_{y_i - h_i \to \infty} f(z_i | y_i) = \lim_{y_i - h_i \to \infty} e^{-A(z_i + B(y_i - h_i)^2 + \frac{1}{2}C^2(y_i - h_i)^2)} = 0$$

2) skew-t model

$$\lim_{y_{i}-h_{i}\to\infty} f(z_{i}|y_{i}) = \lim_{y_{i}-h_{i}\to\infty} \frac{\left[1 + \frac{1}{\nu} \left(\frac{y_{i}-h_{i}}{\sqrt{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}}\right)^{2}\right]^{\frac{\nu+1}{2}}}{\left[\nu + \frac{(y_{i}-h_{i})^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}} + \frac{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}\sigma_{z}^{2}} \left(z_{i} + (y_{i}-h_{i})\frac{\sigma_{z}^{2}}{\sigma_{\epsilon}^{2}+\sigma_{z}^{2}}\right)^{2}\right]^{\frac{\nu+2}{2}}} = 0 \quad \forall z \ge 0$$

E outhing A

biendesse faloreador far bie endicidare anefficience: Le cibi model

ektoria konsiler i hir debowing nëtamoni olor 🗦 p<sup>lat</sup>ire Di er <del>ja k</del>on ektoria e kalendar

. The second of the second of

#### References

- AIGNER, D., LOVELL, C.A.K. AND SCHMIDT, P. (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **6** 21-37
- AZZALINI, A. (1985) A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12 171-178
- AZZALINI, A. AND CAPITANIO, A. (1999) Statistical applications of the multivariate skew normal distribution. *Journal of Royal Statistical Society, B* **61** 579-602
- AZZALINI, A. AND CAPITANIO, A (2002) Distributions generated by permuation of symmetry with emphasis on a multivariate skew t distribution. *Journal of Royal Statistical Society* To appear
- Broeck, J.van den, Koop, G., Osiewalsky, J. and Steel, M. Stochastic frontier models. A Bayesian perspective. *Journal of Econometrics* **61** 273-303
- COELLI, T. PRASADA RAO, D.S. AND BATTESE, G.E. (1998) An Introduction to Efficiency and Productivity Analysis Kluwer Academic Publisher.
- CHRISTENSEN, L.R. AND GREENE W.H. (1976) Economies of scale in U.S. electric power generation. *Journal of Political Economy* 84 655-676
- Greene, W.H (1990) A gamma-distributed stochastic frontier model *Journal of Econometrics* **46** 141-163
- Griffin, J.E. and Steel, M.F.J (2002) Semiparametric Bayesian inference for stochastic frontier models Submitted
- JONDROW, J., LOVELL, C.A.K., MATEROV, I.S., SCHMIDT,P. (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* **19** 233-238
- MEEUSEN, W. AND J. VAN DEN BROECK, J. (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic review* 8 435-444
- RITTER, C. AND SIMAR, L. (1994) Another look at the American elettrical utility data. *Technical report B1348, Institut de Statistique, Universitè Catholique de Louvain* Louvain-la-Neuve (Belgium).
- STEVENSON, R.E. (1980) Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics* **13** 57-66
- TSIONAS, E (2002) Stochastic frontier models with random coefficients. *Journal of Applied Econometrics*? ??-??

UNIVERSITA' DI PADOVA

BIBLIOTECA DI SCIENZE STATISTICHE

VIA C. Battisti, 241 - 35121 PADOVA

Repair of significant

Argyssu Du Boynes, 923, W. Asin Schwindt, P. (1973) Fermich fonskijt estruction of dochreek Production production from teammentleis. Product of New West Steel & School

Akkastant, ja 1985. At dese di diskitations y fant, ne dode jag mound koest. Amenikuum Tutse, ed optikate takat 1779.78

Andrew Color Color Color Color (1995) 15 (1995) Andrew Color Color

er vanden in de de en eeste de l'en de de daaldde leer 20 (2000), ee leadhe e 20 (200). As de eeste Aleis and Digween de eeste Terre de geween and angeleer de leeste en de Communication, place de leeste die ee

Pudate de l'Accepte d'Avarieu (La Cheza alisee) de Acceptante de la companya de l'Accepte mentral. À 2 març de conseptant de la Coura de Communication de La Coura de Coura de Coura de Coura de Coura de Coura d

aterial de la companya de la compan Indignata de la companya de la comp

namen in de la completa de la compansión de la completa de la completa de la completa de la completa de la com La completa de la completa della completa de la completa del completa del completa de la completa del completa del completa della compl

Andreway W. H. (1990) of general combated standards the cost meson. The compared to the expression of the expression of the cost of the co

O brate v. J. R., Alvia (spreau M. R. 1 (2003) Semiger annage javyell in historomo for casha**stic fro**ns Tier escalak Salkerbest

a preparet, and discretion of the contract of Contract in the contract of the speciments contract production than the product of the contract of Contract of Contract of the contract of the

idiodestate, W. Ate. J. ven per Breched. 2. (1977) Etrejene, est motion from CohleDoches production bracedors with compagen error. Andragonal Elementar recess Science 1944.

i Surient, C. obeetinian, Listedd) Amsterniodemeth 'Anne indicated reini-spille, and and remains to Legget 1932 to testgirt de Statisgesco *Universit's Cutheligno de Societ*e Louven Louven universe en (Bildina)

Sign virkskok. M.E. (1936), talediliand functions for generalized so the order energionalist. Tour water ties experies 14 57-66

PSIONAS (2002) Stochastic from her a solds with reading co-flictors. Agreed of Applied Farras-

VOIAS TO ATTERSVINUS

	normal linear model		ALS model		skew-t model	
	mle	95% confidence interval	mle	95% confidence interval	mle	95% interval.
$\beta_0$	-7.2047	(-7.8512 -6.5582)	-7.4071	( -8.0389 -6.7752)	-7.8208	(-8.4419 - 7.1997)
$eta_1$	0.3860	$(0.3126 \ 0.4595)$	0.4081	$(0.3316\ 0.4846)$	0.4549	$(0.3871\ 0.5228)$
$eta_2$	0.0316	$(0.0264 \ 0.0368)$	0.0306	(0.0254, 0.0357)	0.0278	$(0.0231 \ 0.0325)$
$eta_3$	0.2462	$(0.1178 \ 0.3746)$	0.2439	$(0.1186\ 0.3692)$	0.2952	$(0.1724\ 0.4181)$
$eta_4$	0.0792	$(-0.0390\ 0.1974)$	0.0592	$(-0.0606\ 0.1790)$	0.0344	$(-0.0725\ 0.1412)$
$\sigma_z$		<del></del>	0.1558	$(0.0853 \ 0.2844)$	0.0900	$(0.0302\ 0.2682\ )$
$\sigma_\epsilon$	0.1419	(0.1252, 0.1608)	0.1069	$(0.0703\ 0.1627)$	0.0949	$(0.0669\ 0.1347)$
$\nu$		<del></del>		<del>-</del>	4.6461	$(1.912\ 11.2441)$
log lik	65.67		66.14		68.75	

Table 1: Estimates, 95% confidence intervals and maximum loglikelihood values

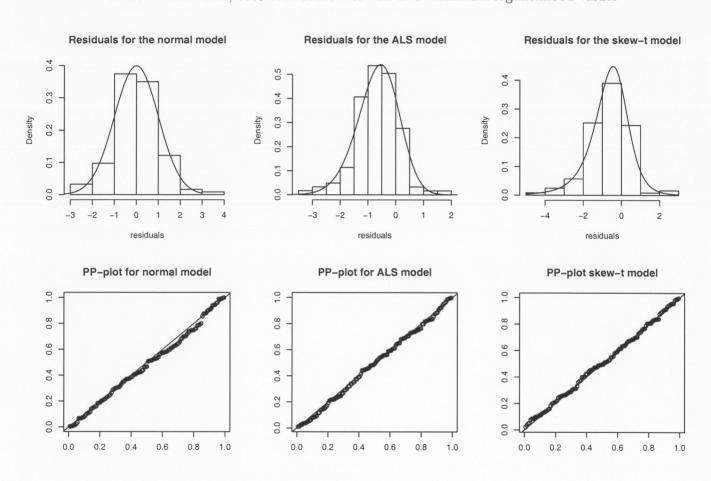


Figure 1: Residuals analysis

	ALS model	skew-t model
$r_1$	0.7325	0.7749
$r_2$	0.9650	0.9408
$r_3$	0.9145	0.9432
$r_4$	0.8980	0.9252
$r_5$	0.9510	0.9524

Table 2: Efficiencies for the first five firms





















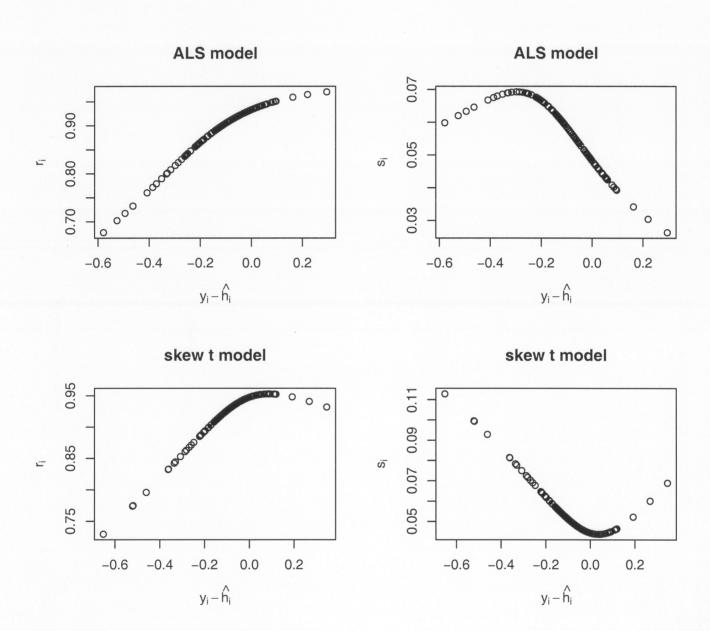


Figure 2: Eefficiencies respect to the estimated residuals